

Library Analytics: Shaping the Future — Applying Data Analysis: Demonstrate Value, Shape Services, and Broaden Information Literacy

by **Rachael Cohen** (Discovery User Experience Librarian, Indiana University Bloomington) <rachcohe@indiana.edu>

and **Angie Thorpe Pusnik** (Digital User Experience Librarian, Indiana University Kokomo) <atthorpe@iuk.edu>

Column Editors: **John McDonald** (EBSCO Information Services) <johnmcdonald@ebSCO.com>

and **Kathleen McEvoy** (EBSCO Information Services) <k.mcevoy@ebSCO.com>

Academic librarians are no strangers to statistics and assessment, and we do not just passively provide access to materials. For decades, librarians have evaluated their collections and programming against changing user needs and interests. In recent years, however, librarians have shifted from assessing their impact primarily through peer comparisons to instead turning inward and assessing their impact against distinctive institutional student success outcomes. With decreasing budgets and increasing calls for evidence of impact, we are having to prove both our value and our expertise at every turn. We are being asked to justify the purchase of expensive tools and journal packages, and we are called upon to demonstrate how our subject expertise directly influences student learning. Rather than waiting to be asked to provide proof of impact, librarians have assumed the driver's seat and begun collecting and analyzing data to assess the value of our collections and demonstrate our importance to the academy.

There is no better time to jump headfirst into proactively illustrating the library's value. Academic librarians grapple with several hard truths every day. For one, we acknowledge but strive to defy the fact that we often only reach and actively consult with a small percentage of our users. The truth remains that a percentage of our users never step foot into the university library. We also contend with the oft-held belief — from both students and faculty alike — that, as “digital natives,” students already know how to search. As we have found, though, search transactions logs often show the exact opposite.

The time is ripe for librarians to find new ways to communicate and work with their users. Failing to do so further removes us from the research process, which reduces our overall value. One powerful solution that may already be available to many librarians is anonymous user data. In particular, anonymous user behavior and search transaction logs from web-scale discovery services offer rich testimony to user skillsets and subject interests. The

authors contend these data are essential toward developing connections with more library users and establishing the pivotal role of the library within the academy.

Case Studies

Both authors work at institutions that launched **EBSCO's** discovery service in fall 2011. By early 2014, largely motivated by the hefty financial investments we were making in discovery, we realized we needed to begin assessing their efficacy. We turned to data both because it was comprehensive, and it documented genuine, unfiltered user behavior. Even the collection of data from searches recommended or conducted by a librarian reflected actual “in the wild” discovery usage. Usage data is, in fact, the record of what actually happened.

We compared user behavior metrics across our two campuses: the flagship **Indiana University** campus in Bloomington, which enrolled nearly 47,000 students in 2013-2014, and the significantly-smaller regional **Indiana University** campus in Kokomo, which enrolled approximately 2,600 students in 2013-2014. We looked at longitudinal usage reports from our discovery vendor and Google Analytics over the three most recent academic years.

Vendor data proved insightful user engagement metrics, such as full text downloads and abstract views. Google Analytics yielded valuable user behavior data, including the devices and browsers used to access discovery, plus the distribution of basic versus advanced user searches. Although we had hypothesized

that user behavior would differ between our campuses, our results mostly complemented one another's, including the fact that — even in 2014 — despite the ubiquitous nature of smartphones, students at both campuses still overwhelmingly used a desktop or laptop computer to access discovery.¹

One year later, in 2015, we returned to our data. Although we had shared our 2014 findings with our colleagues, admittedly, the change we had sought had not yet transpired. In this context, as well as in response to observations that online library resource usage

far exceeded in-person library visits and that university administrators were increasingly interested in overall assessment of student success, we asked ourselves some tough questions:

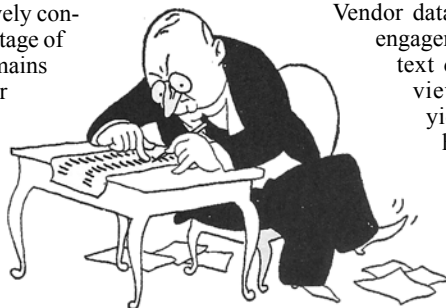
- Discovery services/tools were undoubtedly more prevalent in libraries, but how were students actually searching in them?
- What trends, successes, and challenges do user search queries reveal?
- Moreover, who is even using our discovery services/tools?
- How does user behavior compare across our two institutions: the largest and the smallest campuses in a multi-campus university system?

We also had to ask ourselves how we would even find answers to these questions besides our standard discovery usage data. Focus groups, interviews, and surveys were possibilities, but focus groups and interviews are inherently limited in terms of participation counts, and surveys run the risk of low participation rates. Furthermore, through other studies we have found that students are rather prone to telling you what they think you want to hear, rather than their unfiltered reality. Thus, we instead turned to transaction logs for our dataset. Transaction logs — or user search queries — hold the capability of revealing not only users' information needs but also broader trends and patterns in searching behaviors.

Our methodology involved examining search queries from our respective discovery service logs, which were harvested from Google Analytics. We collected a semester's worth of data, and then we each identified a random sample of 1,677 recorded search queries. To understand who was using discovery at each campus, we then reviewed and categorized each query in our respective datasets, using the Library of Congress (LC) Classification schedule to assign both a class and subclass to each query. Categorizing queries allowed us to pinpoint themes and/or assignments students were working on so that we could then collaborate with our teaching colleagues — both in and outside the library — to further help students succeed in their information-seeking activities.

We initially set out to identify recurrent queries at each campus and then compare results in order to gauge any overarching patterns

continued on page 68



of user behavior across campuses. However, our end project innovated on the initial project plan because we did not simply analyze our search query logs, identify heavy versus light discovery service users, and determine which LC subjects were searched most/least frequently (respectively, social sciences and medicine/world history and agriculture). Rather, we discovered search patterns and themes, which open up unexpected opportunities to engage more deeply with our instruction librarian and teaching faculty colleagues through the application of our search query data analysis. Search patterns included search missteps, such as typographical errors and questions (e.g., “why people travel”). Where our first dive into data analysis provided us with interesting but not necessarily actionable results, our second study on search queries provided practical evidence we can leverage to improve our services.²

Applying Data-Informed Results

We are not the first libraries to use Google Analytics to evaluate our resources, and we acknowledge the appeal of continuing to peel back the layers of collected data. However, rather than just analyzing data, we propose applying data to effect change. Data such as transaction logs allow us to both demonstrate the value of our collections and our expertise to the academy. Librarians with technical and public service responsibilities alike can utilize anonymous usage data to partner with instructors to build students’ information literacy skills, shape library services and resources, and increase overall engagement with users. These activities can, in turn, improve user assignments, perceptions of library services, and overall appreciation for the library.

Collection managers, including subject specialists and liaisons, can use search transactions to improve collection development. Frequently searched titles that are not part of the library’s collection should be considered for acquisition. Titles that are owned may be prime opportunities for outreach regarding print or online library reserves, if allowed. Themes that emerge from subject categorization may also reveal changing disciplinary focal points. For example, if a subject recurs in search queries but is collected at only a minimal or basic level, collection managers may want to explore coverage expansions, in consultation with both faculty and library service providers.

For librarians who serve at public service desks, identifying search patterns and themes can serve as preparation for possible reference questions they may likely encounter during specific semesters. This means that, even if the questions and/or topics are outside their area of expertise, librarians can attune themselves with the optimal resources needed to answer questions. This preparedness, in turn, instills a perception that librarians are truly experts in research and are an important resource to utilize. The librarian becomes as much a resource

to consult as the recommended material in the library’s collection.

Librarians with system responsibilities will likely also find it beneficial to review transaction logs. Log analysis facilitates the identification of similar queries and patterns, which can help ensure librarians learn how users interact with resources. This understanding enables librarians to better advocate for and then implement appropriate new features that assist students in their searching. This knowledge will likely also benefit vendors as they continue to develop intuitive yet effective products for library users.

Last but definitely not least, transaction log analysis assuredly also benefits librarians with instruction responsibilities. Transaction logs reveal common user errors and misunderstandings when searching library resources. Equipped with this knowledge, librarians can proactively advise students regarding how to craft their search queries — including using built-in tools such as auto-suggest keyword features — to avoid problems before they emerge. This is particularly helpful because studies have shown that students are more likely to refine or completely change their search query versus use a facet or click to the second page of results.³ This level of preparedness enables librarians to move away from teaching basic searching skills to instead spend more time covering precise techniques that will enable students to formulate their research questions and match their search strategies to appropriate search tools.

We have spoken in generalities regarding how librarians with different responsibilities may benefit from data or transaction log analysis. Since completing our data analysis, we have begun to translate our results into action. We launched supplementary training sessions for student employees and librarians, recognizing that these groups require different instruction methods and levels of detail. The trainings focused on what is and is not included in our discovery services, discovery facets — such as source types or publications — and content providers, since several search queries included specific facet (e.g., “professional journals about teaching” and “Journal of American Medical Association”) or database names (e.g., “ArtStor” and “factiva”). Our student trainings incorporated independent, hands-on activities, while our librarian trainings involved demonstrations and additional technical explanations.

We are also developing intentional outreach plans to share our results with our teaching faculty. We have identified a limited number of faculty on each campus with whom we would like to initially work. For this pilot phase, we plan to work with faculty who are teaching capstone courses. We will present sample search queries from each faculty member’s discipline and begin a conversation about the instructor’s satisfaction with their students’ coursework. We foresee these conversations then moving to how we can collaborate to build reciprocal value between our faculty and our librarians.

Conclusion

The ACRL Framework for Information Literacy for Higher Education calls librarians to identify core ideas within their own knowledge domain that can extend learning for students, create a new cohesive curriculum for information literacy, and collaborate more extensively with faculty.⁴ The authors uphold data — including transaction logs — as a core idea that affects all of these areas. That is:

- Data empowers librarians to identify where student searches are breaking down so they can improve services and extend learning for students.
- Data analysis and evaluating student searches thus logically extend themselves as integral components of a cohesive curriculum for information literacy initiatives.
- As a component of the information literacy curriculum, sharing data is another strategy for librarians to collaborate more extensively with faculty.

This latter point is significant: On their own, search queries are just data. They lack context, and they are anonymous. However, librarians can use this data as another starting point to have more meaningful conversations with our colleagues; we can extract meaning from our data. This data allows us to connect actual user information needs with our services, with our collections, with our instructors, and with our courses. With this data, we are better equipped than ever before to collaborate with our colleagues, build students’ abilities to formulate appropriate search queries, use information ethically, and meet student success outcomes. 🌱

Endnotes

1. Cohen, Rachael A., and Angie Thorpe. “Discovering user behavior: Applying usage statistics to shape frontline services.” *The Serials Librarian* 69.1 (2015): 29-46. <https://doi.org/10.1080/0361526X.2015.1040194>
2. Cohen, Rachael A., and Angie Thorpe Pusnik. “Measuring query complexity in web-scale discovery: A comparison between two academic libraries.” *Reference & User Services Quarterly* (in press).
3. Examples: Georgas, Helen. “Google vs. the library (part II): Student search patterns and behaviors when using Google and a federated search tool” *portal: Libraries and the Academy* 14.4 (2014): 503-532. <https://doi.org/10.1353/pla.2014.0034>; Brett, Kelsey, Ashley Lierman, and Cherie Turner. “Lessons learned: A Primo usability study.” *Information Technology and Libraries* 35.1 (2016). <https://dx.doi.org/10.6017/ital.v35i1.8965>
4. Association of College and Research Libraries. *Framework for information literacy for higher education*. (2015). <http://www.ala.org/acrl/standards/ilframework>